BERLINER BEITRÄGE ZUR SKANDINAVISTIK

Titel/ title:	*Språk og språkforhold i Sápmi*
Autor(in)/ author:	Michael Riessler/Joshua Wilbur
Kapitel/ chapter:	»Documenting the endangered Kola Saami languages«

In:	Bull, Tove/Kusmenko, Jurij/Rießler, Michael (Hg.): Språk og språkforhold i Sápmi. Berlin: Nordeuropa-Institut, 2007
ISBN:	973-3-932406-26-3
Reihe/ series:	Berliner Beiträge zur Skandinavistik, Bd. 11
ISSN:	0933-4009
Seiten/ pages:	39–82

Diesen Band gibt es weiterhin zu kaufen. This book can still be purchased.

Michael Riessler and Joshua Wilbur

Documenting the endangered Kola Saami languages

1. Introduction

The present paper addresses some practical and a few theoretical issues connected with the linguistic field research being undertaken as part of the *Kola Saami Documentation Project (KSDP)*.

The aims of the paper are as follows: a) to provide general information about the Kola Saami languages und the current state of their documentation; b) to provide general information about KSDP, expected project results and the project work flow; c) to present certain software tools used in our documentation work (Transcriber, Toolbox, Elan) and to discuss other relevant methodological and technical issues; d) to present a preliminary phoneme analysis of Kildin and other Kola Saami languages which serves as a basis for the transcription convention used for the annotation of recorded texts.

1.1. The Kola Saami languages

The Kola Saami Documentation Project aims at documenting the four Saami languages spoken in Russia: Skolt, Akkala, Kildin, and Ter. Genealogically, these four languages belong to two subgroups of the East Saami branch: Akkala and Skolt belong to the Mainland group (together with Inari, which is spoken in Finland), whereas Kildin and Ter form the Peninsula group of East Saami.[1] Thus, from a strictly geographical point of view only Kildin and Ter, spoken on the peninsula, should be regarded as »Kola Saami«. However, the term »Kola Peninsula« is now often used as a synonym for the Russian administrative area (Murmansk District – Ru *Murmanskaja oblast'*) where the four Saami languages are (or were) spoken. From a socio-historical point of view, it makes perfect sense to include the Russian Saami languages in a single group because all four languages share the same contact-linguistic environment and have been exposed to similar assimilation pressure from Russian.

All four Saami languages spoken in Russia are highly endangered due to language shift. Practically all speakers of Kola Saami languages are

[1] Sammallahti 1998, 26–34.

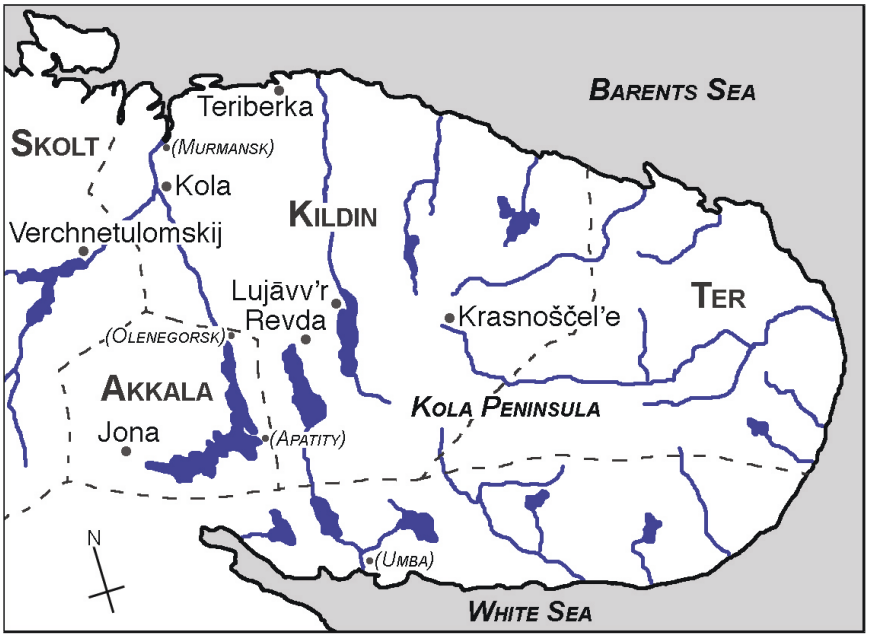


Figure 1: The four original Kola Saami dialect areas and Saami communities near original Saami villages. Towns in parentheses also have Saami inhabitants.

bilingual in Russian and the Kola Saami languages are hardly ever heard in public life today.

Akkala Saami is probably extinct. According to information from different Saami and non-Saami individuals, the last speaker of Akkala passed away in 2003.[2]

Ter Saami is nearly extinct and spoken by perhaps 30 speakers living in different places within and outside of the Murmansk region, such as in Murmansk, Lujāvv'r (Lovozero), Revda, Krasnoščel'e, Umba, and even in St. Petersburg. The mean age of the youngest speakers of Ter is over 50. Note that the number of Ter Saami speakers estimated by KSDP is somewhat higher than the numbers which are normally given in reference books.[3]

---

[2] See also SALMINEN 2007, 235.

[3] For example SERGEJEVA 2002, 107, or SALMINEN 2007, 271, who put the number at only six speakers.

Skolt Saami is a special case. First, with the exception of only a few speakers Skolt Saami is spoken in a relatively compact area by approximately 300 Saami in Finland now.[4] This is perhaps the explanation for the second point, namely that Skolt Saami is better described and probably also somewhat less endangered than Kildin Saami. However, the Notozero dialect of Skolt Saami probably has even fewer speakers in Russia than Ter and is in fact almost extinct. Most of these speakers live close to the original territories of their villages in the northwestern parts of the Murmansk District, mostly in Verchnetulomskij but also in Tuloma, Murmaši and other places.

The number of active speakers of **Kildin Saami** likely amounts to 300. The number of 500–700 speakers, which is found in most reference books[5] seems to be rather optimistic. Still, several older and even a few younger Kildin Saami use their mother tongue in conversation with family members, relatives and friends. On the whole, however, the number of Kildin speakers is decreasing rapidly from year to year and the language must be characterized as severely endangered. Especially among the younger generation, there is a strong decline in active language competence due to the lack of a vibrant speech community and the lack of social motivations for learning and using Saami.

Originally, Kildin was spoken in most parts of the central Kola Peninsula. Today, more or less compact Kildin Saami settlements in or close to their original villages are found only in Lujāvv'r, Revda, Kola, and Teriberka. But small Kildin Saami speach communities are found today in all larger settlements, such as in Murmansk, Olenegorsk, Apatity, etc.

As a result of the forced resettlement of most of the Kola Saami population to Lujāvv'r, this village is nowadays usually regarded as the »Sámi capital« of Russia. Lujāvv'r in fact has by far the densest Saami population today. However, less than one third of the village's approximately 3000 inhabitants are ethnic Saami, and the amount of Saami speakers among these is considerably lower.

The forced centralization of Kola Saami to Lujāvv'r has resulted in

---

4 Ibid., 268.

5 For example Sergejeva 2002, 107, and Salminen 2007, 155 and in the most recent Russian census of 2002 referred to in Scheller 2006, 282; see also Scheller's contribution to this volume.

the mixing of some dialectal features, especially of the different Kildin Saami varieties. Since most Kildin Saami speakers now live in Lujāvv'r, the central and probably most innovative local Kildin Saami variety has been chosen as a normative basis for the language. Consequently, the available teaching materials and dictionaries of Kildin use orthographic and grammatical variants based on this variety. This has led to the innovative Lujāvv'r-dialect of Kildin being usually regarded as the standard variety of Kola Saami.

Of the 6 main dialects of Kildin, at least the four dialects Ārsjogk (Ru *várzinskij dialekt*), Lujāvv'r (Ru *lovozérskij dialekt*), Koarrdegk (Ru *vorón'enskij dialekt*) and Kīllt (Ru *kil'dínskij dialekt*[6]) are still maintained by older speakers, even after their relocation to Lujāvv'r or other places.

1.2. State of linguistic documentation of Kola Saami

The Kola Saami languages are not undocumented; indeed quite a few investigations of certain aspects of their grammar, a few text collections, dictionaries and grammatical descriptions and even some teaching materials are available. Still, lexical and grammatical descriptions are incomplete and there is almost no data available which reflects current language use.

Some of the most important studies of Kola Saami include the descriptive dialect dictionary of T. I. Itkonen,[7] Kert's reference grammar of Kildin,[8] Zajkov's small description of Akkala Saami phonology and morphology,[9] and three dictionaries published by Kuruč, Afanas'eva and Mečkina (Kildin-Russian)[10] Kert (Kildin-Russian-Kildin),[11] and Sammal-

---

6 The name *Kildin*, originating from the name for an island on the Barents Sea coast close to present-day Murmansk, originally referred only to the rather peripheral Kīllt dialect of Kildin Saami. However, *Kildin Saami (language)* (Russian *kil'dinskij (saamskij) jazyk*, North Saami *Gielddasámegiella*) became used as exonym referring to the whole group of neighboring Saami dialects linguistically distinguished from the other Kola Saami dialect groups Akkala, Skolt, and Ter.

7 ITKONEN 1958.

8 KERT 1971.

9 ZAJKOV 1987.

10 KURUČ, AFANAS'EVA and MEČKINA 1985.

11 KERT 1986.

lahti and Chvorostuchina (Kildin-North Saami-Kildin).[12] Text collections of Kola Saami languages are found in T. I. Itkonen 1931, Kert 1961, Szabó 1967; Szabó 1968, E. Itkonen 1985 and Panfilov, Kert and Zajkov 1988. Most of these texts consist of phonologically transcribed and translated stories and fairy tales in Kildin and Ter Saami. Many hours of recorded Kola Saami texts exist in different archives in Russia, Estonia, Finland, Norway and Sweden. Most of the data, however, is inadequately cataloged and not provided with metadata. Annotations are available only for a minor part of these. Currently, none of the archives provides on-line access to the recordings.

2. The Kola Saami Documentation Project (KSDP)

The aim of the Kola Saami Documentation Project is to provide comprehensive linguistic, sociolinguistic and ethnographic documentation of the endangered Saami languages of Russia. Linguistic documentation is focused on Kildin but includes work with the small number of speakers of the other Kola Saami languages in order to produce as complete a survey as possible.

It is the primary goal of the project to systematically record, transcribe, translate and archive the broadest possible variety of spoken language data from the Kola Saami languages. Linguistic documentation, provided with linguistic and ethnographic annotations, is expected to reflect active and passive native-speaker competence, situational and social structuring and the geographical distribution of the Kola Saami languages according to their current patterns of use.

In addition to pure linguistic documentation, Saami language sociology and sociolinguistics are being evaluated with the help of questionnaires and interviews (see Scheller's contribution to this volume for more details on this part of our documentation work).

All data and analyses are saved in and made available through the digital DoBeS archive at the Max Planck Institute for Psycholinguistics in Nijmegen, the Netherlands.

The project is affiliated with the Department for Northern European Studies at Humboldt-Universität zu Berlin and is part of the DoBeS Initiative (Dokumentation bedrohter Sprachen – Documentation of En-

[12] Sammallahti and Chvorostuchina 1991.

dangered Languages, see also http://www.mpi.nl/DOBES) which was launched by the German Volkswagen Foundation in 2000. The time frame for the project is three years and covers September 2005 through August 2008. At present there are two principle researchers, one external principle researcher (from Umeå University), two full time Saami assistants, one Saami student assistant, and one student assistant in Germany working for the project. The project maintains a field office in Lujāvv'r as an important part of its infrastructure. Further information on the current state of research can be found on the project's website: http://www2.hu-berlin.de/ksdp.

2.1. Expected results

The processing of recorded language data is carried out mainly at the Department for Northern European Studies in Berlin. The work in Russia centers on transcription and translation of the recorded data. Members of the Saami community work as project and student assistants as well as main language consultants for the project. Other Saami experts are periodically hired to help with recordings, transcriptions and translations. In this way the project hopes to encourage and to stimulate the passing on of traditional knowledge and language within the Saami community.

The goals of the Kola Saami Documentation Project concern linguistic and sociolinguistic documentation as well as working with the language community. In the framework of documentary linguistics – and as part of the DoBeS program – KSDP considers its first priority to be collecting, annotating, and archiving linguistic data. By the end of 2008, an annotated corpus of texts and sketch grammars for each of these languages shall have been produced. These sketch grammars will be written in both Russian and English and will be made available online from the DoBeS archive.

In addition to this purely linguistic data, sociolinguistic data concerning ethnic Saami and their language habits are currently being gathered as a part of Scheller's PhD project on the revitalization of Kildin Saami. Initial general statistics from this sociolinguistic survey based on 1000 questionnaires will be made available at the end of 2007 or the beginning of 2008.[13]

[13] See also Scheller's contribution to this volume.

The very object of such research, the language itself, is of course crucially dependent on an active community of speakers. In order to support these individuals and their community as well as to promote active use of the endangered Kola Saami languages, the project actively contributes its energies and resources to the language community. Seminars have been held in Lujāvv'r and Murmansk concerning the further development of a common orthography and the language in general. The project organizes language courses, language camps and language speaking circles. School books in Kildin Saami are being produced both with and without Russian translations. The first instance of this is a small anthology of fairy tales collected and written down by Saami teacher Nina Afanas'eva;[14] other collections of other authors are in preparation. With additional funding from the Saami Council, a Russian-Kildin-Russian internet dictionary is being produced; the first version is scheduled to go online at the end of 2007 on the project website. Last but not least, a school grammar book in Russian is due to be released in the summer of 2008. The expected results are summarized as follows:

- Linguistic documentation
  - Annotated text corpus
  - Sketch grammars
- Sociolinguistic survey
- Community language work
  - Seminars on language standardization and development
  - Teaching
  - Kola Saami text editions for schools
  - Russian-Kildin-Russian internet dictionary
  - School grammar of Kildin

Ultimately, the project should result in the building and archiving of an extensive corpus of spoken language to be used for future community language work and for future research. The other points mentioned in

[14] In preparation; to appear in 2007.

the list of expected results remain subordinated to this main goal of the project.

Saami academics and non-academics have legitimately claimed that the policy of previous linguistic research on Kola Saami was to simply use the Saami people as objects of study.[15] Indeed, not very much of the material gathered by different researchers has been given back to the Saami community so far.[16] Although KSDP also uses non-indigenous methodology in order to achieve its goals, our collection and analyses of new data concerning the previously under-studied Kola Saami languages hopefully helps to prepare school grammars, dictionaries, text collections and even descriptive materials that will contribute to a more comprehensive documentation of Kola Saami in the future.

In order to fulfill the wish of the indigenous Kola Saami people to respect cultural and linguistic heritage, researchers like ourselves have been asked to make a contribution to a Kola Saami phonoarchive.[17] However, as long as such an archive does not exist, the storage of the data in a browsable and (at least for the most part) freely accessible multimedia archive (such as the DoBeS archive) along with the publication of some parts of the text corpus in books probably comes closest to the legitimate demand of the community on the researchers to give the gathered data back. Copying the text corpus on DVDs (or other electronic media) for the Saami is an important symbolic gesture; however, access to the data in the DoBeS archive and the publication of texts does more towards guaranteeing the dissemination and long-term persistence of the data.

2.2. Work flow

The following section concentrates on some practical aspects of documentation by describing the work flow used in the project for processing linguistic data from the field.

---

[15] For example SERGEJEVA 2002, 114; see also PORSANGER 2004 for a general discussion on indigenous versus non-indigenous methodology in research.

[16] This is true at least for the huge amount of recorded speech. Still, the three Kildin Saami dictionaries as well as some practical and didactic materials are to some extent also a result of non-Saami linguistic research.

[17] For example by SERGEJEVA 2002, 114.

1. Recording Practically all recordings are done on location in Russia.[18] Most of the recordings are done by one of the primary researchers, frequently with the participation of Saami assistants. In most cases video and audio recordings are captured (stored on a computer) and prepared for transcription while still in the field.

In order to ensure meeting high quality standards for the stored data, we record in uncompressed data formats (wav instead of mp3). Audio recordings are normally done as 16 bit linear PCM files with a 48 kHz sample frequency. If possible, recordings are done using video with an external microphone to provide better sound quality. The microphone is either connected to the camcorder or records to a separate unit. We use the following two mid-range camcorders: a Panasonic NV-GS 400 and a Sony DCR-HC 90. For outdoor recordings the Panasonic camcorder is normally equipped with a RØDE Stereo Video Microphone SVM connected directly to the external microphone slot of the camcorder. Alternatively, a wireless microphone set Sennheiser EW 100 ENG D g2 is used with either a table or a lapel microphone The smaller Sony-camcorder (without external microphone slot) is normally combined with a digital (flash memory card) recorder Edirol R-09. The latter is either used with its built-in microphone or with an external Sony ECM-MS907 table microphone. For other recordings we use a Hi-MD recorder Sony MZ-RH1 with an external microphone Sony ECM-MS907.

2. Raw transcriptions and initial Russian translations First transcriptions and translations of recorded texts are done by the Saami assistants in the project's field office in Lujāvv'r using the program Transcriber (see below section I). A practical Cyrillic orthography based on the existing orthography of Kildin Saami is used for these raw transcriptions. In this, the transcriptions are divided into intonation units as well as speaker and language sections which are time-aligned with the original recordings.

3. Practical phonemic transcription During the next step in the project's work flow, the audio and video recordings (which have been

[18] The one exception so far was in Leipzig, Germany, where recordings with one Kildin speaker were made in the phonetic laboratory at the linguistics department of the Max Planck Institute for Evolutionary Anthropology; these recordings focused specifically on syllable structure, intonation patterns and preaspiration phenomena.

double checked in Transcriber for accuracy) along with the raw transcriptions (including the initial Russian translation) are imported into annotation tiers in the multimedia presentation program ELAN (see below section I); the individual intonation units as well as speaker and language sections retain their time alignment with the recordings. After importing the raw transcription, it is further annotated with a practical phonemic transcription; this is done by primary researchers or student assistants in Germany.

4. Morphemic transcription, glossing, and English translations The phonemic transcription is exported to Toolbox,[19] where additional tiers for morphological make-up, interlinear glossing, and a literal English translation are linked to the transcription. Toolbox is also used to build dictionaries comprising all morphemes found in the text corpus (see below section I).

Morphological annotations and translations with Toolbox are done by primary researchers or student assistants in Germany.

Our interlinear glosses follow the »Leipzig glossing rules«.[20] There are different reasons why we decided to use English as the metalanguage for glosses and grammatical category labels, even though most Saami in Russia do not know English. First of all, English is the most widely used linguistic metalanguage and we ourselves are most familiar with glossing in English. Second, the alternative of preparing Russian grammatical category labels (which are not very commonly used by Russian linguists either) and glosses in Russian is not as beneficial as it may initially seem because this would exclude most non-Russian users (including Saami from outside Russia). Third, the texts are accessible for Kola Saami either as the original transcription (for Saami who have command of the language) or as the literal Russian translation also included. For these reasons, we find the use of English sensible.

5. Finalizing the annotations The finished Toolbox annotation tiers are imported back into ELAN where they are aligned to both the audio

---

[19] In this process, time code information is also exported in order to later reimport the glossed transcriptions into ELAN retaining time alignment.

[20] http://www.eva.mpg.de/lingua/files/morpheme.html.

and video signals as well as to the existing raw transcriptions. Two further annotation tiers for free English and Russian translations are added at this step of the work flow. Additional tiers may include linguistic and anthropological notes. Finalizing the ELAN annotations is mostly done by primary researchers and student assistants in Germany. However, the resulting annotations are also checked by Saami assistants in Russia. In some cases it is necessary to go back to points 2 through 4 in the work flow and apply changes or corrections.

6. Archiving The content of the Kola Saami corpora at the DoBeS archive[21] is organized and updated via the internet using the web-based Language Archive Management and Upload System (LAMUS).[22] LAMUS is used, among other things, to create new corpus structures in the archive and to upload new data resources (video and audio data, pictures, annotations) or metadata to the archive. All archived sessions are provided with metadata; these metadata are divided into the layers listed as follows: Primary researchers or student assistants in Germany are responsible for this step.

1. The session metadata (top node):
   (a) Session Name: a short and clear name for identifying a certain session
   (b) Session Title: a more elaborate identification of the session
   (c) Description: a description of what happens in the session
   (d) Language: which language the description is written in
2. Project: information about the project this particular session belongs to
3. Content:
   (a) Content Type: a classification of the data, the modalities used and the context

[21] The DoBeS archive is part of the Ilse Metadata Description Initiative (IMDI) corpus at the Max Planck Institute in Nijmegen, the Netherlands, and linked to the international Digital Endangered Languages and Musics Archive Network (DELAMAN).
[22] http://www.lat-mpi.eu/tools/lamus.

   (b) Descriptions: a general description of the recorded session
   (c) Languages: a list of all languages used during the session
   (d) Actors: a list of all participants (including the researcher) involved in the session

4. Resources:
   (a) Media Files: a list of the video and sound files that are associated with this session
   (b) Written Resources: references to annotations and transcriptions
   (c) Sources: information about the physical master (e.g. a video tape) which the media files are derived from

5. References: textual references to related literature etc.

3. Linguistic annotation tools

In the following sections, the three main linguistic annotation tools (Transcriber, Toolbox, ELAN) which are of central importance for the documentation work done in KSDP shall be described in detail. However, these are not the only programs which have proven useful for the project; others include Praat (for phonetic analyses) as well as Ukelele and Keyman (for creating keyboard layouts for MacOS and Windows, respectively).[23]

3.1. Transcriber

Transcriber is a relatively user-friendly manual annotation tool. It provides an interface for segmenting long speech recordings (for some common audio formats, among them wav and mp3) into units of any length and adding an aligned transcription to these segments. Furthermore, this

[23] Praat is a computer program with which recorded speech can be analysed, synthesized, and manipulated; it is available for free at http://www.praat.org. The user-friendly unicode keyboard layout editor Ukelele is also free and can be downloaded from http://scripts.sil.org/ukelele. Keyman is commercial keyboard mapping software for Windows.

program allows one to label speaker turns and indicate non-linguistic events, among other things.

Transcriber is distributed as free software under GNU (General Public License) and is available for different platforms (Windows XP, Mac OS X and Linux) online at http://trans.sourceforge.net/. A user manual in English is also available and can be downloaded from the same website.

The main feature of the program is its ability to combine a long audio recording with a text annotation tier. The audio file's wave form is displayed, thus allowing one to easily chunk the text into sequences of any length with the visual aid of the wave form; in this, the text tier is automatically time-encoded into the same chunks. In this way the text annotation units keep their alignment to the original audio signal. Another advantage of working with Transcriber is that segmentation boundaries are displayed both under the wave form of the signal and in the text editor window where they are written. Highlighting in the text editor window is always synchronized with the cursor in the signal window while playing a recording.

Equipping the raw transcription with a Russian translation during the same process of work turned out to be a practical solution for this step. In this, a narrow Russian translation follows the Saami transcription after an asterisk in one and the same segmented unit.

A screen shot of Transcriber can be seen in Figure 2 below.

The only problems we have had with transcriber are of a technical nature (and may be restricted to the version for Mac OS X). Character encoding in Unicode (UTF-8) can be chosen by the user and input in different scripts is possible, at least in principle. However, the program seems to support only the standard Mac OS X keyboards. Typing Cyrillic characters from the Russian keyboard or the Russian phonetic keyboard, for instance, is possible, but only »standard« Cyrillic letters are included in these keyboards. Inputting characters from self-designed keyboard layouts or even from Mac OS X's Character Palette is not possible. As a result, certain Saami Cyrillic characters are not available for use in Transcriber.

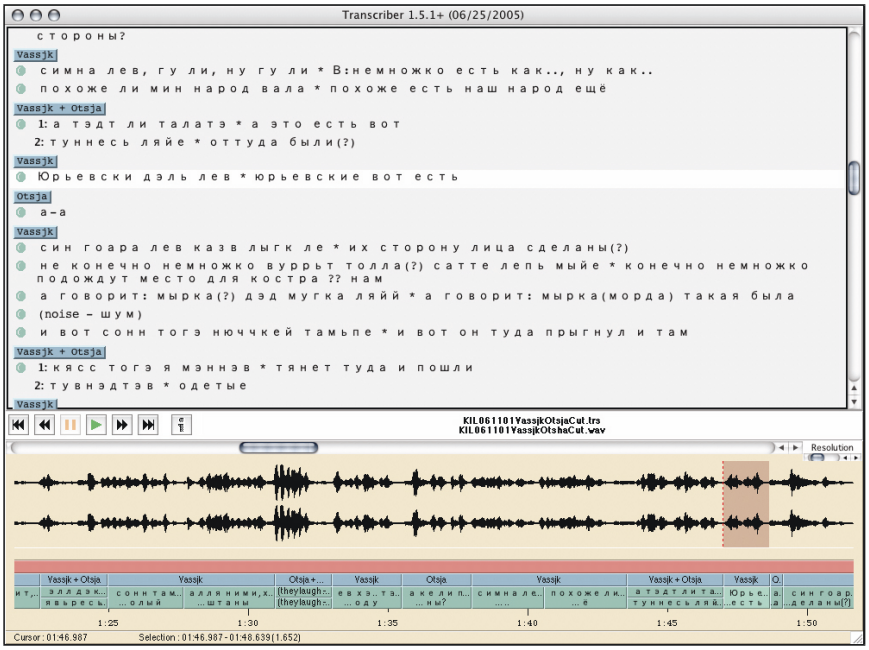


Figure 2: Screen shot of a raw transcript in Transcriber

3.2 ELAN

ELAN is a linguistic annotation tool developed at the Max Planck Institute for Psycholinguistics. It is used to create an unlimited number of complex, hierarchically structured annotations linked to video and audio resources. Essentially, various topic-based layers (called tiers) consist of text describing a particular aspect of the medium, such as sociolinguistic context, phonology, speaker, etc. ELAN links these tiers directly to video and audio recordings, which are both visible as video and waveforms, respectively. The individual text units of such tiers can be time-aligned to the media. Figure 3 shows a transcript imported into ELAN and supplemented with a further transcription tier.

Furthermore, simple corpus searches can be performed on a single document or on a selection of annotation documents. ELAN runs on Windows, Mac OS X and Linux. For further information and to

download the program and an English users manual, see http://www.lat-mpi.eu/tools/elan/.

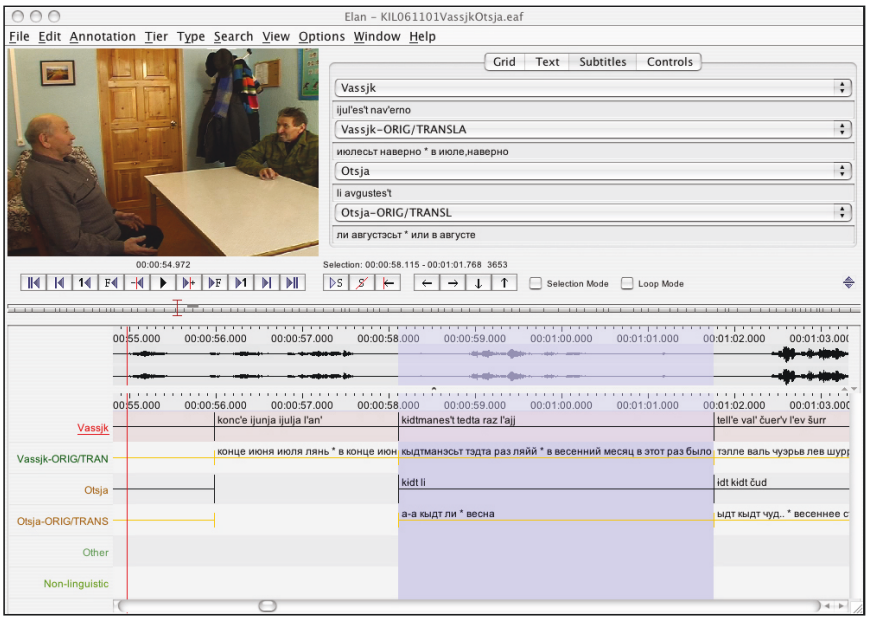


Figure 3: Screen shot of ELAN

3.3 Toolbox

The Toolbox program is essentially a database program tailor-made for linguists. It allows morphological and text corpus data to be organized in familiar and useful linguistic formats like dictionaries, word lists and concordances. A further significant benefit of using Toolbox is that it can automatically perform morpheme analyses of texts based on the data as recorded in the specific project's dictionary database. Toolbox is free software developed by SIL and is available online at http://www.sil.org/computing/toolbox/.

This section aims to look specifically at the application of Toolbox in the documentation of the Kola Saami languages in order to describe what has been done within the context of KSDP and what specific prob-

lems we have encountered so far.

It should be noted that individual Toolbox projects have been created for each of the Kola Saami languages. Each Toolbox project consists of a Toolbox dictionary database file and Toolbox text files, among other things. In the following description of how Toolbox is used for KSDP, our Kildin Toolbox project will be used as an example.

Dictionary Files To begin with, we will look at the Kildin Toolbox dictionary file, which is essentially a list of grammatical and lexical items[24] along with other relevant information about these items. Toolbox comes with its own set of predefined categories arranged into hierarchical tiers which users can adopt just as they are, yet it is flexible because it allows users to alter these tiers as desired and to make up and define new tiers as well. For the Kildin Toolbox project, we have decided to include the information found in Figure 4 below for each entry. These categories reflect the tiers in the Kildin Toolbox dictionary project and are listed in their actual order with their abbreviations shown before their complete descriptions; dots indicate relative position in the tier hierarchy.

The initial and main tier *lexeme* (\lx in Figure 4) includes the lexeme or head word itself which is written in KSDP's orthographic transcription (see below section I) and reflects the word's phonological form. While this tier is called »lexeme«, this terminology is slightly misleading because the tier is also used for all grammatical items such as case suffixes or conjunctions. Tier \a *allomorph* lists dialectal, sociolectal, ad hoc, and other variants of the main lexeme in \lx so that these can also be recognized during text parsing and are included in the main entry they belong to. The next tier, *gradation pattern* (\grad), provides consonant gradation, umlaut and palatalization alternations, if applicable. Tier \u *underlying form* shows the root form which a lexeme belongs to plus any relevant consonant gradation, umlaut or palatalization alternations which cause the current lexeme to differ from its »underlying« root.[25]

---

[24] Generally speaking, a Toolbox dictionary is a collection of a language's individual morphemes. While this may work well for some languages in which morphemes are easily separated from one another, it is not that simple for Saami due to the complexity of its nonlinear morphology.

[25] There are no theoretical assumptions behind our »underlying form«. The reason for introducing this tier is a technical one: it makes it possible to parse the different

Figure 4: Kildin Toolbox dictionary file tiers

\lx	lexeme
. . \a	allomorph
. \grad	gradation pattern
. . \u	underlying form
. . \ps	part of speech
. . \ge	English gloss
. . \gr	Russian gloss
. . \kil	Kildin original
. \dt	date edited
. \nt	notes
. \src	source
. \sln	sociolinguistic notes
. \ety	etymology

The operational stem alternation patterns are presented in parenthesis as part of the morpheme. By clicking on the root shown in this tier, Toolbox automatically switches to the root's entry (a process called *jumping* in Toolbox). The relevant part of speech can be found in the tier *part of speech* (\ps). Tiers \ge *English gloss* and \gr *Russian gloss* provide English and Russian translations (in their respective alphabets), while tier \kil *Kildin original* shows the entry in the original Kildin Cyrillic script.[26] Tier \dt *date edited* indicates the date on which any part of the entry was last modified, added to or deleted. Any relevant additional information which is not found in other tiers is included in tier \nt *notes*; this may include temporary comments for project team members, permanent notes for potential end-users of the dictionary, etc. Tier \src *source* indicates the source of the lexeme, whether this is a written source, a recording, or an elicitation from a certain speaker. Sociolinguistic information (such

---

stem allomorphs as belonging to one head word in the lexicon. See also the discussion on page 62.

[26] For Kildin we use the orthographic variant found in the dictionary by Kuruč et al. 1985 with one slight change: instead of the Latin letter ⟨j⟩ we use the Cyrillic letter ⟨й⟩ (introduced by KURUČ, AFANAS'EVA and VINOGRADOVA 1995) for the voiceless palatal approximant. See also below section I

as a certain entry's dialectal or sociolectal background) as well as etymological information can be found in tiers \sln *sociolinguistic notes* and \ety *etymology*.

The application of this can be understood more easily by looking at an example. Figure 5 shows the entry for the word *kå̄šš̄k* ›cat‹. Here, three asterisks indicate that no information is available in the database. Another example can be seen in Figure 10 on page 63, which is an actual screen shot from Toolbox showing the entry for the word *puaz* ›reindeer‹.

Figure 5: Kildin Toolbox dictionary entry *kå̄šš̄k*

\lx	kå̄ šš̄k
. . \a	***
. \grad	šš̄k - šk
. . \u	kå̄šš̄k
. . \ps	n
. . \ge	cat
. . \gr	ко́шка
. . \kil	коашшк
. \dt	24/May/2006
. \nt	***
. \src	NA
. \sln	dial. Arsjogk
. \ety	< Ru кошка

Text files and morpheme analyses A very important feature of Toolbox is its ability to perform morphological parsing of texts. These texts can be simple words or expressions from field notes, transcribed recordings, or even literary or other texts written previously. Each text forms its own »database« in that it is entered as a file of its own. If a text has not already been divided into intonation units in Transcriber (see above section I), as is the case with literary texts, then Toolbox can automatically break the text into in smaller, more easily manageable chunks based on punctuation. At any rate, these divisions make the text easier to work with on a computer screen and to refer to when using a concordance.

For the Kildin Toolbox project, we include the information found in Figure 6 below for all units of a text. The category names reflect the tier names in the Kildin Toolbox text project and are listed in the actual order with their abbreviations shown before their complete descriptions; dots indicate relative position in the tier hierarchy.

Figure 6: Kildin Toolbox text file tiers

\id	text identification
\block	ELANBlock
. \ref	reference
. . \ELANBegin	timecode start
. . \ELANEnd	timecode end
. . \ELANParticipant	participant
. \txc	text cyrillic
. \tx	text
. . \mph	morphemes
. . \ps	part of speech
. . \ge	English gloss
. \lt	literal translation
. \otr	original translation Russian
. \nt	notes

The initial tier *text identification* (\id in Figure 6) is simply the name of the text and occurs once at the very beginning of the text file. Tiers \block *ELANBlock*[27] and \ref *reference* allot a reference number to each individual unit for use with concordances and word lists. The following tiers \ELANBegin *timecode start*, \ELANEnd *timecode end* and \ELANParticipant *participant* all contain information about time code alignment and the speaker involved for sharing with ELAN[28]. The actual text in Cyrillic and in the project's orthographic Roman alphabet translit-

[27] This tier is only present if the text is the transcription of a recording processed using Transcriber and ELAN (see above sections I and I). The same is true for the other ELAN tiers.

[28] Specifically, this information originates from ELAN, is imported into Toolbox for parsing, then exported back into ELAN for the final thorough annotation of a recording.

eration can be found in tiers \txc *text cyrillic* and \tx *text*, respectively. The latter text tier forms the basis for interlinear morpheme glossing and corresponds to the dictionary entries in the dictionary database. Tier \mph *morphemes* shows the results of the automatic morpheme analysis done by Toolbox (described below), while tiers \ps *part of speech* and \ge *English gloss* are copied from the individual dictionary entries of the respective parsed morphemes. Tier \lt *literal translation* provides a more or less word-for-word English translation of the Kildin text, just as the original word-for-word Russian translation provided by the Saami assistant(s) who originally transcribed the text can be found in tier \otr *original translation Russian*. Finally, tier \nt *notes* allows any meta-commentary to be added; this can include any relevant information, such as questions about or notes on the present grammatical annotation. Three asterisks again indicate that the information is irrelevant or missing in the dictionary file.

Once a text has been entered, Toolbox can automatically perform a morpheme analysis (parsing). In doing so, Toolbox works through the text from beginning to end of the tier *text* (\tx) looking for matching entries at both the lexeme level and the allomorph level of the dictionary entries described in section I above. When a text is first entered into Toolbox, the *morpheme*, *part of speech* and *English gloss* tiers are not visible because Toolbox provides these automatically during parsing. This can be seen in the example in Figure 7 below. Here, the sentence *Mēhkal tar'm vus'te lūz, nu, sūll' lūz* ›Michael bought salmon today, well, salt-salmon‹, which is taken from a longer text called »Bad Fish« has been entered into Toolbox, but no morpheme analysis has been performed.

By clicking the interlinearize button, Toolbox automatically conducts a morpheme analysis. If multiple matching items are found in the dictionary due to homophony[29], the user is asked to select the correct entry from a list of possibilities; Toolbox then enters the chosen morpheme. The program can also be »taught« by way of formulas to automatically make certain choices based on the morphosyntactic environment of a homophone so that, for instance, if the suffix *-e* occurs following a lex-

[29] Of course, for Toolbox this is not really homophony, but a matter of homographic entries.

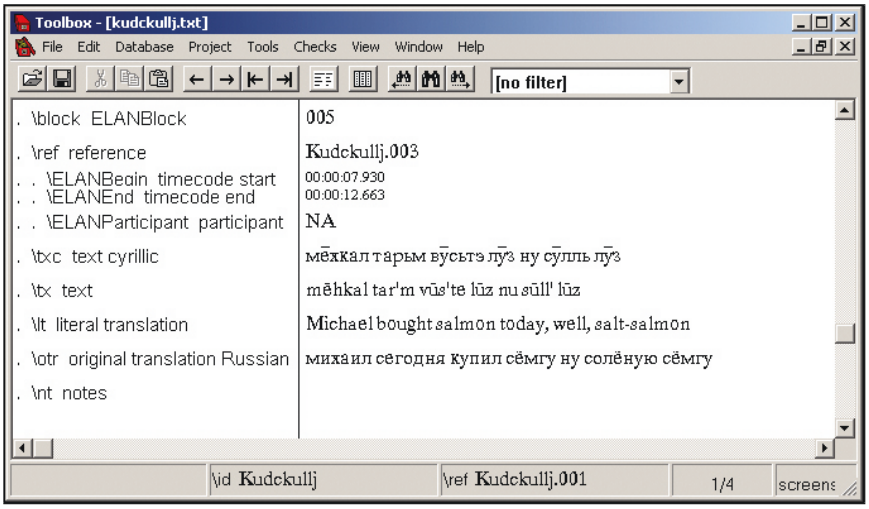


Figure 7: Screen shot of Kildin Toolbox text entry before parsing

eme item listing as a noun, Toolbox automatically knows that this can neither be the infinitive suffix *-e* nor the third person singular past suffix *-e*, but possibly the illative singular suffix *-e*. In case Toolbox does not find any matching entries in the dictionary file, the user is given the opportunity to create a new entry in the dictionary database based on the unidentified word or morpheme. It is therefore easy to increase the total volume of dictionary entries by working through new texts and constantly updating the dictionary by adding new words encountered in the texts.

After a morpheme has successfully been identified, it is automatically entered in the morpheme tier of the text file under the appropriate word in the text tier, while the respective part of speech and English gloss are also added in alignment. The completed morpheme analysis of the example sentence found in Figure 7 can be seen in Figure 8 below; now the morpheme, part of speech and English gloss tiers are also visible.
Once a large amount of data has been collected, comprehensive dictionaries can be produced containing all or only some of the information included in the different tiers of the Toolbox dictionary file. Such dictio-

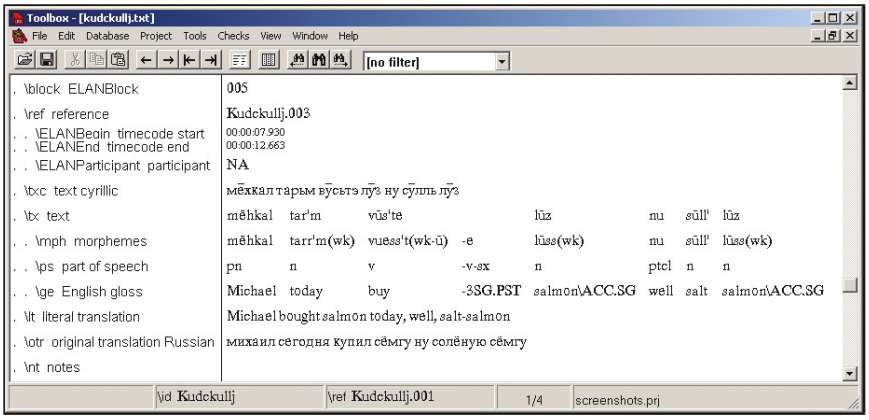


Figure 8: Screen shot of a Kildin Toolbox text entry after successful parsing

naries shall be printed or otherwise made accessible to the speech community. In addition, word lists and concordances based on the entire text corpus can also be created by Toolbox.

Kildin Saami Complications While the basic application of Toolbox as described above may sound complicated but essentially effective for dealing with language data, the reality of using it for Saami languages is unfortunately rather problematic. Technical problems on the one hand and linguistic problems on the other hand complicate the situation.

Toolbox is most efficient when dealing with languages which have strictly linear morphology. The abundance of nonlinear morphology in Saami languages in general and Kola Saami in particular make using Toolbox especially challenging because it is frequently impossible to draw one-to-one relationships between meanings (both grammatical and lexical) and morphemes due to widespread homophony among inflectional affixes (as, e.g., *e* which can be glossed as -INF, -1SG.PST, -3SG.PST, -IMP.PL, -GEN.SG, -GEN.PL, -ACC.SG, -PART, -PRED, NEG=). A special challenge is also posed by the various stem mutations caused by consonant gradation, umlaut and palatalization. For instance, to return to the example in Figure 5, *kå̄šš̄k* is the word for ›cat‹ in the singular, while the plural *kå̄šk* is marked exclusively by the presence of the weak form of

consonant gradation for this word, and is indeed the shortening of the consonant cluster /ʃːk/ to /ʃk/. Because the allomorph \a *kå̄šk* is included in the entry \lx *kå̄ššk*, Toolbox recognizes it as the word meaning ›cat‹; however, a problem occurs because it does not alone mean ›cat‹, but in fact ›cat\PL‹. This fact needs to be recorded in both the dictionary file and in the interlinear glossing of the text file. The situation becomes more complicated because weak consonant gradation can also indicate genitive singular: *kå̄šk* ›cat\GEN.SG‹ and the weak form also occurs as a stem allomorph in several inflected forms of the word (like in *kå̄šk-en'* ›with the cat‹).

For Toolbox to recognize each of these possibilities as potential glosses and then parse the right morpheme appropriately, each would have to have its own individual entry in the dictionary file indicating its particular nonlinear meaning, resulting in an unnecessarily large and somewhat misleading dictionary. Instead, all different allomorphic forms of a root should ideally be included under one single dictionary entry. Our solution is to enter allomorphic variants into Toolbox in subsections called *sense entries* (abbreviated as \se in the dictionary file) which include the particular allomorph and information about it stem alternations, part of speech and an English gloss. In this, each allomorph is recorded in its own sense entry located under a single lexical item (\lx) in the hierarchy. The different glosses for each allomorph are separated by semicolons in the English gloss tier.

Figure 9: Extract from dictionary entry *kå̄ššk* with allomorphs

```
\lx      kåššk
. . \u   kåššk
. . \ps  n
. . \ge  cat-; cat(NOM.SG)
. \se    kåšk
. . \u   kåššk(wk)
. . \ps  n
. . \ge  cat-; cat\NOM.PL; cat\GEN.SG
```

Figure 9 illustrates the situation with the example entry *kå̄ššk*, which has two allomorphs due to the strong and weak forms of consonant grada-

tion. Together, these allomorphs have 5 different grammatical meanings. The initial glosses in both subentries ›cat-‹ are followed by a hyphen and the second meaning of the strong stem allomorph is glossed as NOM.SG. However, only one gloss ›cat‹ is necessary in the actual Toolbox dictionary entry since nominative singular is an unmarked case in Kildin and the hyphen occurs as part of the suffixes in the dictionary entries and while parsing with Toolbox. A screen shot of the Toolbox dictionary entry for *puaz* ›reindeer‹ (including its subentries) is found in figure 10.

There are no theoretical assumptions behind our »underlying form«. The decision to posit such forms for all dictionary entries was made for practical reasons in order to be able to work more efficiently with Toolbox. The »derived« stem allomorphs compose the subentries, and their glosses are in the respective subentry tiers. We normally use the elicitation form (the NOM.SG form of nouns and the infinitive stem of verbs) for the underlying form. Note that whereas the »main stem allomorph« of the example *kå̄šš̄k* listed in the \lx tier is in the strong grade, the morpheme used for the main lexical entry of the example *puaz* shows the weak grade.[30]

Two more examples for the interlinear glossing for a few stem allomorphs of the words ›cat‹ and ›reindeer‹ can be seen in example 1.

(1) Glossing of different stem allomorphs

a.
\tx	kå̄šš̄k	kå̄šš̄k-en’	kå̄šk	kå̄šk-en’
\mph	kå̄šš̄k	kå̄šš̄k-en’	kå̄šš̄k(WK)	kå̄šš̄k(WK)-en’
\lit	cat	cat-ESS	cat\GEN.SG	cat-COM.SG

b.
\tx	puaz	pūʒ-en’	pūʒ-en’	
\mph	puaz	pūʒ(STR)-ESS	pūʒ(STR-ū)-COM.SG	
\lit	reindeer	reindeer-ESS	reindeer-COM.SG	

The situation and solution is the same when dealing with umlauts and other kinds of nonlinear morphology and with allomorphy. However, this may not be the best solution because of a further problem with how Toolbox deals with homophony. Specifically, Toolbox automatically returns all possible glosses from the entire lexical entry (and not just from the subentry) for the user to choose the correct parse from when it en-

[30] In fact, the strong stem *pūʒ-* is the »underlying form« from a diachronic point of view.

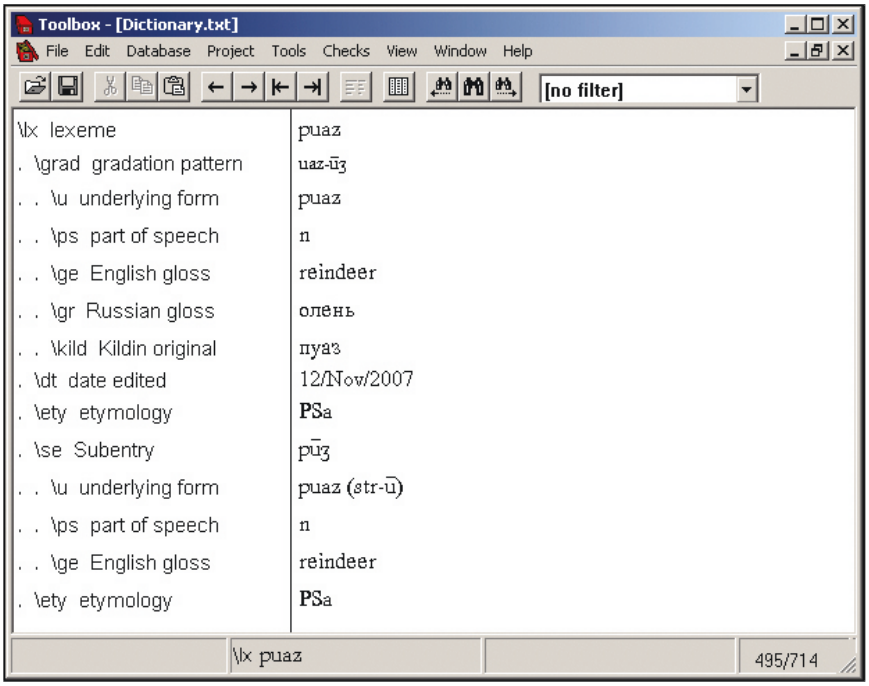


Figure 10: Screen shot of the Kildin Toolbox dictionary entry *puaz* ›reindeer‹

counters a morpheme with multiple matches during parsing. Each time this happens, the user is confronted with what is frequently a long list of very similar suggested parses made up of every potential combination of all the morphemes recognized and their glosses. It can be very time consuming to search for the correct combination, even though the user most likely already knows how the word should parse.

A further disadvantage to using Toolbox is of a more technical nature and specific to Macintosh computers. To date, Toolbox runs only on Windows operating systems. There is a very limited alpha version for Mac available, while full versions for Mac and Linux are in development, but may never be completed due to complications and the ability to run the Windows version of Toolbox using virtual machine software. Despite these challenges, Toolbox remains a valuable tool in documenting the Kola Saami languages.

4. KSDP's phonemic transcription and its phonological background

The main goal of KSDP is documentation and archiving in order to make the data available to and useful for community based language development initiatives, as well as for further research. As opposed to all extant Kola Saami text collections which have transcriptions based solely on the interpretation of the individual linguists, KSDP annotations shall be made available together with the original audio files. This has two main advantages: on the one hand, precise phonological or even phonetic annotations are not required; on the other hand, other researchers shall be able to improve on our analyses or add further annotation tiers. However, to a certain extent grammatical analyses are essential to the actual documentation process. Our grammatical analyses thus shape our documentation work in vital ways. For instance, the phoneme inventory determined by KSDP forms the basis for the practical transcription used for writing Kola Saami, while the project's compilation of morphemes is used as the basis for morpheme glossing. The following sections go into detail in describing our practical phonemic transcription convention and present Kola Saami vowel and consonant systems.

4.1 A practical transcription for Kildin

The practical phonemic transcription used by KSDP can be considered a compromise between a broad phonological (or even phonetic) transcription, as traditionally used in Uralistics, and a true orthographic transcription in either the existing Cyrillic orthography or a transliteration of it.

Using a phonological or even phonetic transcription would be a time consuming task and is in fact not imperative since all annotations will be directly linked to the original recording and future researchers interested in phonetic details can thus easily augment the annotations with phonetically narrower transcriptions.

The existing Kildin Saami orthography, on the other hand, is no alternative either since it is not very transparent. At least in some respects it is not at all faithful to the phonological system of the language. In addition, the Kildin Saami orthography can hardly be considered established since three different orthographical variants exist. The main differences

between them are the representation of preaspiration either as ⟨xx⟩, as ⟨h⟩, or as ⟨'⟩ (apostrophe) and of the voiceless palatal approximant either as ⟨йх⟩,[31] as ⟨j⟩, or as ⟨й̦⟩. Even though only a few Saami seem to use the orthography, each different orthographic variant has its own group of adherents, and they frequently have very emotional views concerning the other variants.

As long as the Kola Saami do not come to an agreement about one orthographic variant on their own, we will provide the annotations of the unpublished parts of the text corpus only with a Latin based phonemic transcription. For published parts of the text corpus, as well as for a school grammar, we have initially decided to use the orthographic variant found in the large Kildin-Russian dictionary.[32] This dictionary not only constitutes the largest normative piece of work on Kildin Saami and can thus be considered a reference work, but the orthographic variant used in it we find to be more consistent than the orthography used in the smaller school dictionary.[33] We would however suggest applying one slight change to the orthography of Kuruč et al. 1985: Instead of the Roman letter ⟨j⟩ we use the Cyrillic letter ⟨й̦⟩[34] to denote the voiceless palatal glide. By doing this, all voiceless sonorants are marked consistently with a tail on the letter for the respective voiced sonorant.[35]

On the other hand we are of the opinion that the Kildin Saami writing system can be used for the other Kola Saami languages with a few adjustments for each language. Because the phonological systems of the Kola Saami languages are not very different, all Kola Saami could use one and the same Cyrillic orthography with differences merely occurring in lexical items. Nonetheless, it is necessary that the speakers themselves agree on a single orthography.

As a basis for our practical phonemic transcription we essentially use a transliteration of the existing Cyrillic orthography for Kildin. For cer-

---

[31] According to Kert's phonological analysis, however, the voiceless palatal approximant (/j̥/) does not exist in Kildin (consider, for example Kert 2005, 2).

[32] Kuruč, Afanas'eva and Mečkina 1985.

[33] Kert 1986.

[34] Introduced by Kuruč 1995.

[35] However, for Kola Saami text editions which are published through KSDP but not authored by ourselves, we leave the choice of orthographic variant to the individual Saami authors.

tain phonemes which are not adequately represented in our assessment, we use additional characters or character combinations. Most of them are taken from the Latin based orthography developed for Skolt Saami in Finland.[36]

In principle it would be possible to transliterate our Roman transcription back to a Cyrillic script, rendering an orthography which is more or less identical to the the already existing one. The character chart in the appendix of this paper presents the existing Cyrillic script of Kildin compared to the single characters of KSDP's practical phonemic transcription, and a somewhat adapted version of the Cyrillic script which could be used for all Kola Saami languages.

Some major differences between KSDP's practical phonemic transcription and the existing Cyrillic orthography of Kildin can be found in the representation of the voiceless sonorants ⟨м̨, н̨, й̨, р̨, л̨⟩ as *mh, nh, jh, rh, lh* and the representation of palatal, palatalized, and »half-palatalized« consonants. According to our phonological analyzis (see section I, especially page 71 below) there is no such feature as »half-palatalization« and the »half-palatalized« coronal plosives ⟨тЬ, дЬ⟩ and the »half-palatalized« nasal ⟨нЬ⟩ are better analyzed as palatalized *t', d', n'*. What is written as a »fully palatalized« coronal nasal in the orthography ⟨нь⟩, on the other hand, has to be analyzed as palatal nasal [ɲ] (at least in the speech of some older Kildin speakers) and is represented as *nj* in KSDP's practical phonemic transcription. Similar to the latter, the palatal lateral approximant [ʎ] (in our transcription *lj*) is actually represented as combination of characters ⟨льй⟩ in the Cyrillic orthography as well.

Another difference is the double writing of the voiced plosive geminates. These are represented as consonant clusters ⟨бп, дт, гк⟩ in the Kildin Saami orthography. The reason for this is that the long consonant is perceived as being strongly devoiced towards the end of its articulation. In this respect Kildin Saami does not behave differently from other Saami languages in which the voiced geminate stops are represented orthographically doubled *bb, dd, gg*.

There are also some differences between KSDP's practical phonemic transcription and the Skolt Saami orthography. These mainly regard the

---

[36] Compare, for example, SAMMALLAHTI and MOSNIKOFF 1991.

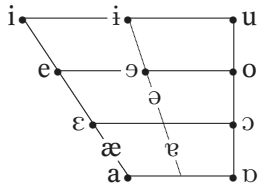
location of the palatalization sign and the representation of preaspiration. Skolt Saami is normally analyzed as having suprasegmental palatalization. The apostrophe, marking palatalization, is subsequently written before the consonant (or between the vowel and consonant center). Regardless whether palatalization in Kildin and the other Kola Saami languages has to be analyzed as a segmental feature or a suprasegmental feature, we decided to follow the existing Kildin Saami orthography and mark palatalization either at the end of a single consonant or geminate or after the first consonant of a consonant cluster.

The (historical) preaspiration of voiceless geminate stops and affricates is marked with *h* (the IPA character for the voiceless glottal fricative) in KSDP's phonemic transcription. This is done disregarding the actual synchronic phonological status of these sounds in the Kola Saami languages (or varieties) in question.

4.2 Vowels and diphthongs

Figure 11 presents our preliminary analysis of vowel phones found in the Kola Saami languages.

Figure 11: Kola Saami vowel phones



Since only part of these vowels are to be analyzed as phonemes in the different Kola Saami languages and some of them are subject to certain positional restrictions (for example the short central vowels /ə/ and /ɐ/ which only occur in unstressed syllables in Kildin Saami), a practical phonemic transcription of the Kola Saami language will manage with a smaller amount of vowel characters.

As in the Kildin and Skolt Saami orthographies, KSDP's transcription uses digraphs for writing phonemic diphthongs: i.e. /ue, ia, ea/ etc.

Vowel length, which is distinctive in all Kola Saami languages, is marked with a macron, i.e.: /ā, ē/. This is different from the Skolt Saami orthography where long vowels are written with the double vowel character. The macron, however, is taken from the Kildin Saami orthography and could also be used in case diphthongs have to be marked for phonemic length as well: /ōā/.

Kildin Saami vowel phonemes Our preliminary analysis of the vowel system of Kildin Saami, which is more or less similar to the one presented by Kert,[37] yields the following inventory of vowels and diphthongs: /i, ɨ, u, e, o, a, å, ue, ua, ie, ea/. Length is distinctive in stressed syllables, except for the high central vowel /ɨ/ and the diphthongs.

Table 1: Kildin Saami vowel phonemes in KSDP's orthographical transcription

i ī	ɨ	u ū
e ē		o ō
a ā		å å̄

4.3 Consonants

One Kola Saami characteristic – as compared to the western Saami languages – is the relatively large consonant inventory, which is mostly due to the fact that almost all consonants have the phonemic oppositions voiced-unvoiced, long-short, and palatalized-non-palatalized.[38] Table 2 lists all phones which we have identified so far in the different Kola Saami languages (length distinctions are not marked).

There are two phonological features in the Kola Saami languages which deserve special attention: palatalization and preaspiration.

Palatalization A process of palatalization causes the secondary modification of consonants with a non-palatal primary place of articulation: the primary articulation is accompanied by the raising of the tongue dor-

[37] KERT 2005, 2.

[38] Note, however, that palatalization has also been analyzed as a suprasegmental feature, at least for Skolt Saami.

sum against the hard palate. Palatalization thus has to be distinguished from **palatal articulation**, which indicates the primary articulation of a consonant. The place of articulation in the latter is the hard palate, the articulation organ is the tongue dorsum.

Palatalization occurs as a reflex of an apocoped etymological front vowel (PSa **i* or **e*) in the Kola Saami languages. Other Saami languages do not have phonologically palatalized consonants. Palatal consonants, however, are a remnant of Proto-Saami found in all Saami languages. The innovative palatalization in Kola Saami thus resulted in a phonological system with both palatals and palatalized consonants. As a result of this development phonological oppositions arose such as the palatalized postalveolar nasal /nʲ/ with the palatal nasal /ɲ/ and of the palatalized postalveolar lateral approximant /lʲ/ with the palatal lateral approximant /ʎ/, see example 2.

(2) Semi-minimal triplets of plain, palatalized, and palatal consonants in Kildin

a. ⟨мāнн⟩ *mānn* /maːnː/ ›moon, month‹
⟨маннь⟩ *mann'* /manːʲ/ ›egg‹
⟨маннь⟩ *mannj* /maɲː/ ›daughter-in-law (of a woman)‹

b. ⟨пāлл⟩ *pāll* /paːlː/ ›ball‹
⟨малль⟩ *mall'* /malːʲ/ ›juice‹
⟨мāлльй⟩ *māllj* /maːʎː/ ›rust‹

The existence of such a phonological opposition between palatal and palatalized consonants is rarely attested crosslinguistically[39] and a change towards merger or separation of the two phonemes would not be unexpected.

Indeed, the palatalized-palatal oppositions are subject to change and seem to be absent at least from the speech of younger Kildin Saami. Interestingly, this change is taking place in different directions. Whereas the nasals /nʲ/ and /ɲ/ are often merged to /nʲ/, the palatal lateral approximant /ʎ/ is often articulated as a consonant cluster [lj] and in this way becoming phonemic by being distinguished from its opponent, the palatalized dental approximant /lʲ/.

---

[39] Compare, for example, Stadnik 2002.

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Alveolar-palatal	Palatal	Velar	Glottal
Plosive	p b pʲ bʲ			t d tʲ dʲ			c ɟ	k g kʲ gʲ	
Nasal	m m̥ mʲ m̥ʲ			n n̥ nʲ n̥ʲ			ɲ	ŋ ŋʲ	
Trill				r r̥ rʲ r̥ʲ					
Fricative	 ß	f v fʲ vʲ	ð	s z sʲ zʲ	ʃ ʒ ʃʲ ʒʲ		ç	x ɣ xʲ	h hʲ
Affricate				ts dz tsʲ dzʲ		tʃ dʒ			
Approximant							j j̥		
Lateral approximant				lʸ l̥ʸ lʲ l̥ʲ			ʎ		

Table 2: Kola Saami consonant phones

The second development is in fact reflected by the orthographic realization of the palatal sound as consonant cluster: ⟨льй⟩. In the case of the nasal the orthography represents the palatalized sound as »half-palatalized« but the palatal sound as »full-palatalized: ⟨нѣ⟩ vs. ⟨нь⟩.

Kert describes »half-palatalization« as being characteristic for the coronal plosives as well.[40] However, in this case minimal pairs are available only for plain /d/, /dː/ and /t/ with their »half-palatalized« counterparts ⟨дѣ⟩ /dˑ/, ⟨дтѣ⟩ /dtˑ/ and ⟨тѣ⟩ /tˑ/.[41] The latter are in fact better analyzed as palatalized /dʲ/ and /tʲ/.

Jurij Kusmenko (p.c.) suggested that the notion of »half-palatalization« is probably the result of a missinterpretation inferred by the Russian philological background of the researchers. The articulatory difference between palatalized /dʲ/ and /tʲ/ in Saami and Russian is quite obvious since in Saami these are true palatalized sounds [dʲ], [tʲ] whereas in Russian the sounds are clearly affricated [dʑ], [tɕ] (note their almost palatal place of articulation). In this sense the /dʲ/ and /tʲ/ in Russian are more »fully« palatalized articulated than in Saami.

Regarding the nasal, the notion of »half-palatalization« and »full-palatalization« was inferred here as well by analogy and the palatal nasal has been perceived as the »more« palatalized half of this pair. However, this terminology does not help much in describing the actual oppositions found in the Saami phonological system.

Another controversy regarding palatalization caused by the apocope in Kola Saami is whether or not it should be analyzed as a segmental or suprasegmental feature. Finnish scholars describe palatalization in Skolt Saami as suprasegmental; examples can be seen in 3. In Skolt Saami orthography suprasegmental palatalization is marked with an apostrophe between a vowel and the following consonant center.

(3) Suprasegmental palatalization in Skolt Saami[42] [maːnː] /maːnː/ ›moon, month‹
mââ'nn [mɐ̟ːnːʲ] /maːʲnː/ ›egg‹
mââ'nnj [mɐ̟ːɲː] /maːʲɲː/ ›daughter-in-law (of a woman)‹

---

[40] Kert 2005, 2.
[41] The dot /ˑ/ marks »half-palatalization« in Kert's transcription.
[42] The example words are taken from Sammallahti and Mosnikoff 1991, transcription is our own.

There are several reasons to analyze palatalization as suprasegmental in Skolt. First of all, palatalization of a consonant stem triggers phonetic effects on both the consonant(s) and the neighboring vowels: especially the stressed vowel preceding the consonant center becomes somewhat advanced in its articulation while the stem consonant(s) is (are) palatalized. Furthermore, consonant cluster palatalization is more perceivable on the first consonant than it is on the second one.

In the other Kola Saami languages, palatalization also tends to be articulated by moving leftward from the end of the consonant center (i.e. away from the right word-edge where the triggering vowel apocope took place historically). As in Skolt the preceding vowel shows advanced articulation and in a consonant cluster palatalization shows up on the first consonant rather than on the second one. At least in Kildin, however, palatalization also occurs on word-initial consonants as well, even though this is restricted to Russian loanwords. Consider, for example, the semi-minimal pair *lānnt* ›puddle‹ and *l'ānnht* < Ru *lénta* ›band; tape‹.[43]

Since the advanced articulation of a vowel in the environment of a palatalized consonant seems to be a natural surface phonetic feature (which occurs in other languages as well) it would also be possible to analyse palatalized consonants as segments. At least for Kildin this analysis is preferable. Otherwise we would end up with two distinct palatalizations: one which is suprasegmental and the other which is segmental and restricted to word-initial consonants.

Preaspiration A period of voicelessness at the end of the vowel, nasal, or liquid preceding the onset of the closure of voiceless stops or affricates (in some languages even fricatives) is typically known as preaspiration.

---

[43] In orthographical spelling palatalized word-initial consonants occur frequently in Kildin Saami, for example кӣлл transliterated as *k'ī̆ll* ›language‹, пӗннэ̄ transliterated as *p'ē̆nne* ›dog‹, and сӕм transliterated as *s'ā̆m* ›beard‹. According to our analysis the palatalization is not phonological in these examples, which are better transcribed as /kīll/, /pienne/, /seam/. The orthographical spelling with monophthongs perhaps reflects an ongoing change (in which the monophthongization of the diphthongs /ie/ and /ea/ would in fact result in the phonologization of word-initial palatalization). Or the orthography simply reflects a misinterpretation of the Kildin Saami phonology because of a comparison with the (diphthong-less) Russian language.

(4) Preaspiration and consonant stem gradation in North, Skolt, and Kildin Saami in the word ›sheath‹

NSa	*dohppa*	[hːp]	/tohːpa/	NOM.SG	– *dohpa*	[hp]	/tohpa/	GEN.SG
SkSa	*topp*	[ʰpː]	/top/	NOM.SG	– *toopp*	[ʰp]	/toːp/	GEN.SG
KSa	*tōhp*	[hp]	/toːhp/	NOM.SG	– *tōp*	[p]	/toːp/	GEN.SG

Preaspiration is normally described as being characteristic of most Saami languages. In example I above, however, true preaspiration is only found in Skolt Saami where voiceless plosives and affricates are articulated with preaspiration. In the North Saami example there is no period of voicelessness before the plosive but instead a short or long glottal fricative sound.

The cognate sound is realized differently among speakers of Kola Saami languages, either as [ʰpː] with true preaspiration, as [hp] with a glottal fricative, or even as [xp] with a velar fricative.

What historically was preaspiration of voiceless plosives and affricates (i.e., a feature of a consonant) has thus developed into a fricative sound (i.e., its own consonantal segment) in North and in (at least some variants of) Kola Saami. However, whereas the cluster [hp] in the North Saami example I could be analyzed as sequence of two single phonemes (the first one of which can either be long or short), the cluster is problably phonemic in the Kola Saami languages.

Kildin Saami consonant phonemes Our preliminary analysis of the consonant system of Kildin Saami differs in some respect from those presented by Kert.[44]

As described on page 72 we analyze palatalization in Kildin Saami as a segmental feature and not as suprasegmental. Regarding the palatals we adopt the phonological system of more conservative variants in which the palatal nasal and the palatal lateral approximant are still existent.

In the case of preaspiration, on the other hand, Kildin Saami's phonological system is more innovative than that of Skolt, for example. True (phonetic) preaspiration is lacking in Kildin Saami and the etymological voiceless plosive and affricate geminates are articulated as clusters of a

---

[44] Kert 2005; see also other descriptions of Kildin by Kert and other Russian scholars.

p hp b bb p’ hp’ b’ bb’	t ht d dd t’ ht’ d’ dd’		k hk g gg k’ hk’ g’ gg’
mh mmh m mm mh’ mmh’ m’ mm’	nh nnh n nn nh’ nnh’ n’ nn’	nj nnj	ŋ ŋŋ ŋ’ ŋŋ’
	rh rrh r rr rh’ rrh’ r’ rr’		
f ff v vv f’ ff’ v’ vv’	s ss z š šš ž s’ ss’ z’ š’ šš’ ž’		x xx x’ xx’
	c hc ʒ ʒʒ c’ hc’ ʒ’ ʒʒ’	č hč ǯ ǯǯ	
		jh jjh j jj	
	lh llh l ll lh’ llh’ l’ ll’	lj llj	

Table2: Kildin Saami consonant phonemes in KSDP’s orthographic transcription

fricative and a single voiceless stop or affricate. This is true at least for most speakers who articulate either a voiceless glottal fricative /h/ (as in North Saami) or even a voiceless velar fricative /x/. The latter articulation might be the result of Russian influence. The question remains, however, if these are clusters of two single consonant phonemes or if the cluster as such is phonemic.

The best evidence for arguing that (etymological) preaspiration is a segment of its own in Kildin can be found in the fact that the fricative sound can be palatalized and is then realized as voiceless palatal fricative [ç]. Consider the following examples with and without palatalized consonant clusters: ***nāh’p*** /naːhʲp/ [nạːçp] ›cup‹ (represented in the orthography either as ⟨нāххьп⟩, ⟨нāhпь⟩, or ⟨нā’пь⟩) and ***māhte***

/maːhte/ [maːhte] ›to be able to‹ (represented in the orthography either as ⟨māxxтэ⟩, ⟨māhтэ⟩, or ⟨mā'тэ⟩).

Another argument would be that these clusters resyllabify between the fricative sound and the stop, as in /naːhʲ.p-e/ [cup-PART]. However, voiced geminates also resyllabify in the middle of the sound and can hardly be regarded as clusters of two single phonemes, consider /kuːlʲ.lʲ-e/ [fish-PART]. On the other hand, morpheme boundaries are never shifted along with resyllabification, which would probably count as an argument for the phonemic unity of the cluster.

Because of these assumptions we have decided to analyse /hp, ht, hk, hc, hč/ as phonemic clusters. In stem gradation these consonant clusters alternate with the respective voiceless single consonant: ***nāh'p – nāp'*** (represented in the orthography either as ⟨nāxxьп – nāпь⟩, ⟨nāhпь – nāпь⟩, or ⟨nā'пь – nāпь⟩), ***māht- – māt-*** (represented in the orthography either as ⟨māxxт- – māт-⟩, ⟨māhт- – māт-⟩, or ⟨mā'т- – māт-⟩).

5. Character chart

In the character chart on pages 76 through 80 the existing Kildin Saami Cyrillic graphemes are listed with the corresponding Roman graphemes used in KSDP's practical phonemic transcription as well as Cyrillic graphemes which could be used for an extended Kola Saami orthography. Note that some of the listed Kildin Saami Cyrillic letters are used only in one of the three orthographic variants. We also provide the Unicode and the corresponding IPA character for all graphemes.

5.1 Further explanations of the character chart

In the character chart, *n.a.* means that no equivalent grapheme is necessary, either because the Kildin letter simultanously denotes a certain vowel and the palatalization of the preceeding consonant (the Cyrillic letters ⟨е, ё, и, ю, я⟩) or it denotes the non-palatalization of a syllable-final consonant (the Cyrillic letter ⟨ъ⟩). Graphemes are written in paranthesis if two graphemes describe one single sound (e.g. ⟨льй⟩ or *lj* for /ʎ/). Note also that we suggest using the grapheme combinations ⟨йе, йо, йи, йу, йа⟩ word-initially (instead of ⟨е, ё, и, ю, я⟩).

Existing Kildin Saami Orthography (Cyrillic)		KSDP's Orthographical Transcription (Roman)		Proposed Kola Saami Orthography (Cyrillic)		Explanation	
Letter	Unicode	Letter	Unicode	Letter	Unicode	IPA	Note
А	0410	A	0041	А	0410	/a/	
а	0430	a	0061	а	0430		
Ӓ	04D2						1)
ӓ	04D3						
Ā	0410 + 0304	Ā	0100	Ā	0410 + 0304	/a:/	2)
ā	0430 + 0304	ā	0101	ā	0430 + 0304		
		Â	00C2	Â	0410 + 0302	/ɐ/	3)
		â	00E2	â	0430 + 0302		
Б	0411	B	0042	Б	0411	/b/	
б	0431	b	0062	б	0431		
В	0412	V	0056	В	0412	/v/	
в	0432	v	0076	в	0432		
Г	0413	G	0047	Г	0413	/g/	
г	0433	g	0067	г	0433		
		Ǧ	01E6	Ґ́	0403	/ɟ/	
		ǧ	01E7	ґ́	0453		
Д	0414	D	0044	Д	0414	/d/	
д	0434	d	0064	д	0434		
(ДЖ)		Ǯ	01EE	(ДЖ)		/ʤ/	
(дж)		ǯ	01EF	(дж)			
(ДЗ)		Ʒ	01B7	(ДЗ)		/z/	
(дз)		ʒ	0292	(дз)			
Е	0415	*n.a.*		Е	0415	/ʲe/	
е	0435			е	0435		
Е	0415	(JE)		(ЙЕ)		/je/	
е	0435	(je)		(йе)			
Ё	0401	*n.a.*		Ё	0401	/ʲo/	
ё	0451			ё	0451		
Ё	0401	(JO)		(ЙО)		/jo/	
ё	0451	(jo)		(йо)			
Ё̄	0401 + 0304	*n.a.*		Ё̄	0401 + 0304	/ʲo:/	2)
ё̄	0451 + 0304			ё̄	0451 + 0304		
Ё̄	0401 + 0304	(JŌ)		(ЙŌ)		/jo:/	2)
ё̄	0451 + 0304	(jō)		(йō)			
Ē	0112	*n.a.*		Ē	0112	/ʲe:/	
ē	0113			ē	0113		
Ē	0112	(JĒ)		(ЙĒ)		/je:/	
ē	0113	(jē)		(йē)			
		Ê	00CA	Ê	0415 + 0302	/ə/	3)
		ê	00EA	ê	0435 + 0302		

Existing Kildin Saami Orthography (Cyrillic)		KSDP's Orthographical Transcription (Roman)		Proposed Kola Saami Orthography (Cyrillic)		Explanation	
Letter	Unicode	Letter	Unicode	Letter	Unicode	IPA	Note
Ж	0416	Ž	017D	Ж	0416	/ʒ/	
ж	0436	ž	017E	ж	0436		
З	0417	Z	005A	З	0417	/z/	
з	0437	z	007A	з	0437		
ҳ	04BA	H	0046	ҳ	04BA	/ʰ/, /h/	
h	04BB	h	0048	h	04BB		
И	0418	I	0049	И	0418	/ʲi/, /i/	
и	0438	i	0069	и	0438		
И	0418	(JI)		(ЙИ)		/ji/	
и	0438	(ji)		(йи)			
Й	04E2	Ī	012A	Й	04E2	/ʲi:/	
й	04E3	ī	012B	й	04E3		
Й	04E2	(JĪ)		(ЙЙ)		/ji:/	
й	04E3	(jī)		(йй)			
Ӆ	0419	J	004A	Ӆ	0419	/j/	
й	0439	j	006A	й	0439		
Ј	0408					/j̥/	4)
ј	0458						
Ӈ	048A	(JH)		Ӈ	048A	/j̥/	
ӈ	048B	(jh)		ӈ	048B		
К	041A	K	004B	К	041A	/k/	
к	043A	k	006B	к	043A		
		Ǩ	01E8	Ќ	040C	/c/	
		ǩ	01E9	ќ	045C		
Л	041B	L	004C	Л	041B	/l/	
л	043B	l	006C	л	043B		
Ӆ	04C5	(LH)		Ӆ	04C5	/l̥/	
ӆ	04C6	(lh)		ӆ	04C6		
(ЛЬЙ)		(LJ)		Љ	0459	/ʎ/	
(льй)		(lj)		љ	0409		
М	041C	M	004D	М	041C	/m/	
м	043C	m	006D	м	043C		
Ӎ	04CD	(MH)		Ӎ	04CD	/m̥/	
ӎ	04CE	(mh)		ӎ	04CE		
Н	041D	N	004E	Н	041D	/n/	
н	043D	n	006E	н	043D		
Ӊ	04C9	(NH)		Ӊ	04C9	/n̥/	
ӊ	04CA	(nh)		ӊ	04CA		

Existing Kildin Saami Orthography (Cyrillic)		KSDP's Orthographical Transcription (Roman)		Proposed Kola Saami Orthography (Cyrillic)		Explanation	
Letter	Unicode	Letter	Unicode	Letter	Unicode	IPA	Note
Ӈ	04C7	Ŋ	014A	Ӈ	04C7	/ŋ/	
ӈ	04C8	ŋ	014B	ӈ	04C8		
(НЬ)		(NJ)		Ӊ	040A	/ɲ/	
(нь)		(nj)		ӊ	045A		
О	041E	O	004F	О	041E	/o/	
о	043E	o	006F	о	043E		
Ӧ	04E6						5)
ӧ	04E7						
�originally Ō	041E + 0304	Ō	014C	�originally Ō	041E + 0304	/o:/	2)
ō	043E + 0304	ō	014D	ō	043E + 0304		
		Õ	00D5	Õ	041E + 0342	/ɔ/	6)
		õ	00F5	õ	043E + 0342		
		Ȭ	022C	Ȭ	041E + 0342 + 0304	/ɔ:/	6)
		ō̃	022D	ō̃	043E + 0342 + 0304		
(ОА)		Å	00C5	Å	0410 + 030A	/ɒ/	7)
(оа)		å	00E5	å	0430 + 030A		
(ŌĀ)		Å̄	00C5 + 0304	Å̄	0410 + 030A + 0304	/ɒ:/	7)
(ōā)		å̄	00E5 + 0304	å̄	0430 + 030A + 0304		
П	041F	P	0050	П	041F	/p/	
п	043F	p	0070	п	043F		
Р	0420	R	0052	Р	0420	/r/	
р	0440	r	0072	р	0440		
Ҏ	048E	(RH)		Ҏ	048E	/r̥/	
ҏ	048F	(rh)		ҏ	048F		
С	0421	S	0053	С	0421	/s/	
с	0441	s	0073	с	0441		
Т	0422	T	0054	Т	0422	/t/	
т	0442	t	0074	т	0442		
У	0423	U	0055	У	0423	/u/	
у	0443	u	0075	у	0443		
Ӱ	04F0						5)
ӱ	04F1						
Ӯ	04EE	Ū	016A	Ӯ	04EE	/u:/	
ӯ	04EF	ū	016B	ӯ	04EF		
Ф	0424	F	0046	Ф	0424	/f/	
ф	0444	f	0066	ф	0444		
Х	0425	X	0058	Х	0425	/x/	
х	0445	x	0078	х	0445		
Ц	0426	C	0043	Ц	0426	/ts/	
ц	0446	c	0063	ц	0446		

Existing Kildin Saami Orthography (Cyrillic)		KSDP's Orthographical Transcription (Roman)		Proposed Kola Saami Orthography (Cyrillic)		Explanation	
Letter	Unicode	Letter	Unicode	Letter	Unicode	IPA	Note
Ч	0427	Č	010C	Ч	0427	/ʧ/	
ч	0447	č	010D	ч	0447		
Ш	0428	Š	0160	Ш	0428	/ʃ/	
ш	0448	š	0161	ш	0448		
Ъ	042A	*n.a.*		Ъ	042A	*n.a.*	
ъ	044A			ъ	044A		
Ы	042B	Ɨ	0197	Ы	042B	/ɨ/	
ы	044B	ɨ	0268	ы	044B		
Ӹ	04F8						4)
ӹ	04F9						
Ы	042B + 0304					/ɨː/	2)
ы	044B + 0304						
Ь	042C	’	02BC	Ь	042C	/ʲ/	
ь	044C			ь	044C		
Ь	048C						8)
ь	048D						
Э	042D	E	0045	Э	042D	/e/	
э	044D	e	0065	э	044D		
Ӭ	04EC						1)
ӭ	04ED						
Э	042D + 0304	Ē	0112	Э	042D + 0304	/eː/	2)
э	044D + 0304	ē	0113	э	044D + 0304		
Ю	042E	*n.a.*		Ю	042E	/ʲu/	
ю	044E			ю	044E		
Ю	042E	(JU)		(ЙЮ)		/ju/	
ю	044E	(ju)		(йю)			
Ю	042E + 0304	*n.a.*		Ю	042E + 0304	/ʲuː/	2)
ю	044E + 0304			ю	044E + 0304		
Ю	042E + 0304	(JŪ)		(ЙЮ)		/juː/	2)
ю	044E + 0304	(jū)		(йю)			
Я	042F	*n.a.*		Я	042F	/ʲa/	
я	044F			я	044F		
Я	042F	(JA)		(ЙЯ)		/ja/	
я	044F	(ja)		(йя)			
Я	042F + 0304	*n.a.*		Я	042F + 0304	/ʲaː/	2)
я	044F + 0304			я	044F + 0304		
Я	042F + 0304	(JĀ)		(ЙЯ)		/jaː/	2)
я	044F + 0304	(jā)		(йя)			
’	02BC					/ʰ/	9)

1) The letter denotes »half-palatalization« of a consonant preceeding the respective vowel, see also above page 71.
2) The Cyrillic letters A, E, YA, YERU, YU, YO do not exist as precomposed letters with a macron in Unicode and have to be written as combinations of the respective Cyrillic letter and the COMBINING MACRON.
3) The Cyrillic letters A and E do not exist as precomposed letter with a CIRCUMFLEX in Unicode and have to be written as combinations of the respective Cyrillic letter and the COMBINING CIRCUMFLEX.
4) The Cyrillic letter JE denotes the voiceless palatal approximant only in the second version of the Kildin Saami orthography.[45]
5) The letter denotes »half-palatalization« of a consonant preceeding the respective vowel only in the first version of the Kildin Saami orthography.[46]
6) The Cyrillic letter O does not exist as a precomposed letter with a TILDE in Unicode and has to be written as a combination of the Cyrillic letter O and the COMBINING TILDE; for an additional macron the COMBINING MACRON has to be used.
7) The Cyrillic letter A does not exist as precomposed letter with the RING ABOVE in Unicode and has to be written as combination of the Cyrillic letter A and the COMBINING RING ABOVE; for an additional macron the COMBINING MACRON has to be used.
8) The letter denotes »half-palatalization«, see also above page 71.
9) The apostrophe is used as the letter for preaspiration only in the third version of the Kildin Saami orthography.[47]

6. Acknowledgements

Our special thanks appreciation are deserved by all Saami which have been or are working with us so far, especially our project assistants Nadežda Zolotuchina, Nina Šaršina, Svetlana Danilova (all from Lujāvv'r), our main language consultants Nina Afanas'eva, Zoja Gerasimova (both from Murmansk), Zoja Nosova (from Verchnetulomskij), Aleksandra Antonova, Genadij Lukin, Anastasija Mozolevskaja, Ekata-

45 As in KURUČ, AFANAS'EVA and MEČKINA 1985.
46 As in KERT 1986.
47 As in KURUČ, AFANAS'EVA and VINOGRADOVA 1995.

rina Korkina and the brothers Vasilij (Vass'k), Evdokim (Otsja) and Petr Galkin (all from Lujāvv'r), Anastasija Nielsen (from Neiden), as well as our student assistant Anna Afanas'eva (from Murmansk).

Other team members have provided valuable advice on earlier drafts of this paper, particularly Jurij Kusmenko, who drew our attention to certain aspects relevant to our phonological analysis, and Elisabeth Scheller. The usual disclaimers apply.

Earlier fieldwork on Kola Saami carried out by M. Rießler was supported by the *Gesellschaft für bedrohte Sprachen e.V.*.

References

Itkonen, Toivo Immanuel: *Koltan- ja kuolanlapin sanakirja = Wörterbuch des Kolta- und Kolalappischen. 1–2*. (= Lexica Societatis Fenno-Ugricae; 15), Helsinki: Suomalais-ugrilainen seura, 1958.

Kert, Georgij M.: *Obrazcy saamskoj reči: materialy po jazyku i fol'kloru saamov Kol'skogo poluostrova (kil'dinskij i iokan'gskij dialekty)*. Moskva: Nauka, 1961.

Kert, Georgij M.: *Saamskij jazyk (kil'dinskij dialekt). Fonetika, morfologija, sintaksis*. Leningrad: Nauka, 1971.

Kert, Georgij M.: *Slovar' saamsko-russkij i russko-saamskij*. Leningrad: Prosveščenie, 1986.

Kert, Georgij M.: »Saamskij jazyk«. In: Vinogradov, Viktor Aleksejevič (ed.): *Jazyki Rossijskoj Federacii i sosednix gosudarstv. Enciklopedija v trjox tomax*. Vol. 3. Moskva: Nauka, 2005. 1–8.

Kuruč, Rimma D., Nina E. Afanas'eva and Iraida V. Vinogradova: *Pravila orfografii i punktuacii saamskogo jazyka*. Murmansk: Institut Jazykoznanija Rossijskoj Akademii Nauk Murmanskij Sektor Lingvističeskich Problem Finno-Ugorskich Narodnostej Krajnego Severa, 1995.

Kuruč, Rimma D., Nina E. Afanas'eva and Ekatarina I. Mečkina: *Saamsko-russkij slovar' = Sām'-rūšš soagknehk'*. Moskva: Russkij jazyk, 1985.

Panfilov, Vladimir Z., Georgij M. Kert and Petr M. Zajkov: *Obrazcy saamskoj reči*. Petrozavodsk: Karelskij filial AN SSSR, 1988.

Porsanger, Jelena: »An essay about indigenous methodology«. In: Theodorsen, Per Bäckström; Kristin Jernsletten; Cathrine (ed.): *Spe-*

cial issue on Northern minorities. (= Nordlit; 15). [Tromsø]: Faculty of Humanities, University of Tromsø, 2004. 105–120.

SALMINEN, Tapani. »Europe and North Asia«. In: *Encyclopedia of the world's endangered languages*. London – New York: Routledge, 2007. 211–280.

SAMMALLAHTI, Pekka: *The Saami Languages. An Introduction*. Kárášjohka: Davvi Girji, 1998.

SAMMALLAHTI, Pekka and Anastasija CHVOROSTUCHINA: *Unna sámi-sām' saam'-sámi sátnegirjjáš = Udc' sámi-sām' saam'-sámi soagknegka*. Ohcejohka: Girjegiisá, 1991.

SAMMALLAHTI, Pekka and Jouni MOSNIKOFF: *Suomi-koltansaame sanakirja = Lää'dd-sää'm sää'nnke'rjj*. Ohcejohka: Girjegiisá, 1991.

SCHELLER, Elisabeth: »Die Sprachsituation der Saami in Russland«. In: HORNSCHEIDT, Antje, Kristina KOTCHEVA, Tomas MILOSCH and Michael RIESSLER (eds.): *Grenzgänger. Festschrift zum 65. Geburtstag von Jurij Kusmenko*. (= Berliner Beiträge zur Skandinavistik; 9). Berlin: Nordeuropa-Institut der Humboldt-Universität, 2006. 280–290.

SERGEJEVA, Jelena: »The Eastern Sámi languages and language preservation«. In: *Samiska i ett nytt årtusende*. (= ANP; 2002:717). Nordiska ministerrådet, 2002. 107–116.

STADNIK, Elena: *Die Palatalisierung in den Sprachen Europas und Asiens. Eine areal-typologische Untersuchung*. (= Tübinger Beiträge zur Linguistik; 461). Tübingen: Gunter Narr Verlag, 2002.

SZABÓ, László: *Kolalappische Volksdichtung (aus den Dialekten in Kildin und Ter). Gesammelt und herausgegeben von László Szabó*. (= Abhandlungen der Akademie der Wissenschaften in Göttingen. Philologisch-historische Klasse; 3:68). Göttingen: Vandenhoeck & Ruprecht, 1967.

SZABÓ, László: *Kolalappische Volksdichtung (aus den Dialekten in Kildin und Ter). Zweiter Teil nebst grammatischen Aufzeichnungen. Gesammelt und herausgegeben von László Szabó*. (= Abhandlungen der Akademie der Wissenschaften in Göttingen; 72). Göttingen: Vandenhoeck & Ruprecht, 1968.

ZAJKOV, Petr M.: *Babinskij dialekt saamskogo jazyka. Fonologo-morfologičeskoje issledovanije*. Petrozavodsk: Karelija, 1987.